Class 3: Sample Spaces & Axioms of Probability

Tuesday, September 10, 2019 11:39 AM

- HW 1 (optional) due Thurs, 1pm, Gradescope
- Office hours: 12:50-1:50 Sept 11 / 2:15 - 3:15 Sept 18 onward } wednesdays, 306 Mudd

---

(ch 1 Bertsekas-Tsitsiklis)

Recap:

- **Sample Space**: Ω = set of all possible outcomes

  Ex: $\Omega = \{H, T\}$ (coin), $\Omega = \{1, 2, \dots, 6\}$ (die), $\Omega = \{A+, A, A-, B+, \dots\}$ (grade)

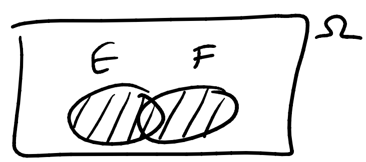
- **Event**: a subset of the sample space, $E \subseteq \Omega$

  Ex: "event that you get an A" = $\{A+, A, A-\}$

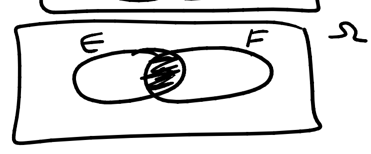
  → **Implication**: We say an event E implies an event F if $E \subseteq F$. This means every outcome of E is also in F.

  Ex: E = "get an A+" = $\{A+\}$, F = "get an A +/-" = $\{A+, A, A-\}$ } $E \subseteq F$

Operations:



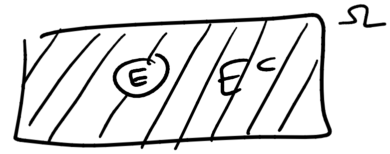
- Union $E \cup F$ = "E or F or both"



- Intersection $E \cap F$ = "E and F"
- Complement E^c = "not E"

- Intersection $E \cap F$ = "E and F"

- Complement E^c = "not E"



Laws:

- $E \cap F = F \cap E$
- $E \cap (F \cap G) = (E \cap F) \cap G = E \cap F \cap G$

(same for union)

- Distributive: $E \cap (F \cup G) = (E \cap F) \cup (E \cap G)$

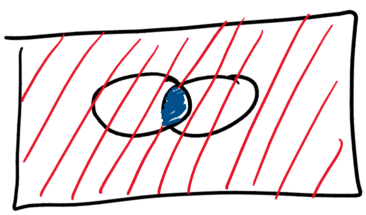
  $E \cup (F \cap G) = (E \cup F) \cap (E \cup G)$

- De Morgan's Laws:

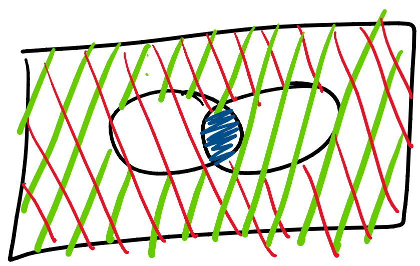
$$(E \cap F)^c = E^c \cup F^c$$

"not both E and F" = "not E or not F"

$$(E \cup F)^c = E^c \cap F^c$$



$E \cap F$ $(E \cap F)^c$



E^c F^c

$E^c \cup F^c$ = everything but blue region

Ex: $\Omega = \{1,2,3,4,5,6\}$, E = "even" $= \{2,4,6\}$

F = "no more than 2" $= \{1,2\}$

E^c = "odd" $= \{1,3,5\}$, $F^c = \{3,4,5,6\}$,

$E^c \cap F^c = \{3,5\}$

$E \cup F = \{1,2,4,6\} \rightarrow (E \cup F)^c = \{3,5\}$

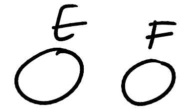
$\rightarrow E^c \cap F^c = (E \cup F)^c$

$$\to \quad E^c \cap F^c = (E \cup F)^c$$

Definitions:

- Two events E and F are disjoint if $E \cap F = \emptyset$.

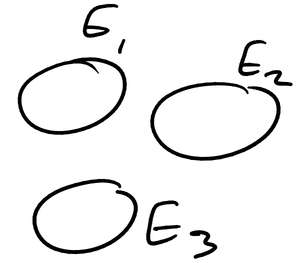
  "At most one of these events can happen, not both."



- A collection of events $E_1, E_2, \ldots$ are mutually exclusive if $E_i \cap E_j = \emptyset$ for all i, j with $i \neq j$.

  "At most one of the E_i's can happen."

  - Same as disjoint for two events.



Ex: $\Omega = \{HH, HT, TH, TT\}$ 2 coin flips

E = "both coins are the same" $= \{HH, TT\}$

F = "at least one heads" $= \{HH, HT, TH\}$

G = "both tails" $= \{TT\}$

$\to$ F and G are disjoint because $F \cap G = \emptyset$, they have no outcomes in common.

$\to$ $E \cap F = \{HH\} \neq \emptyset$, $E \cap G = \{TT\} \neq \emptyset$

- $E_1 = \{HH, HT\}$, $E_2 = \{TH\}$, $E_3 = \{TT\}$ are mutually exclusive.

Note: E and E^c are always disjoint. $E \cap E^c = \emptyset$.

Note: E and E^c are always disjoint. $E \cap E^c = \emptyset$.

Axioms of Probability

(or P)

A probability measure on a sample space Ω is a function $\mathbb{P}$ which assigns to each event a real number ("probability"), satisfying the following properties:

① (Non-negativity) $\mathbb{P}(E) \geq 0$ for all events E.

② (Normalization) $\mathbb{P}(\Omega) = 1$.

③ (Additivity) For two disjoint events E and F,

$$\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F).$$

Consequences:

- $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$

  Proof: Recall E and E^c are disjoint. So axioms ③ and ② give

  $$\mathbb{P}(E) + \mathbb{P}(E^c) \underset{③}{=} \mathbb{P}(E \cup E^c) \underset{E \cup E^c = \Omega}{=} \mathbb{P}(\Omega) \underset{②}{=} 1.$$

- $\mathbb{P}(\emptyset) = 0$. Apply above with $E = \Omega$, $\mathbb{P}(\emptyset) = 1 - \mathbb{P}(\Omega) = 1 - 1 = 0$.

- More general version of ③:

  If $E_1, E_2, \ldots, E_n$ are mutually exclusive, then

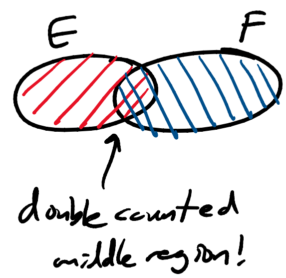
  $$\mathbb{P}(E_1 \cup \ldots \cup E_n) = \mathbb{P}(E_1) + \ldots + \mathbb{P}(E_n).$$

or $\mathbb{P}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \mathbb{P}(E_i)$.

Also works if $n=\infty$: $\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i)$.

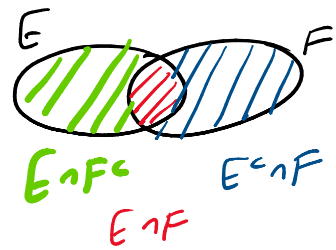
- For <u>any</u> two events E and F, not necessarily disjoint,

$$\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F).$$



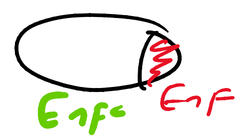
<u>Proof</u>:

The events $E \cap F^c$, $E \cap F$, and $E^c \cap F$ are mutually exclusive, and their union is $E \cup F$.



Axiom ③ implies

$$\begin{aligned} \mathbb{P}(E \cup F) = {} & \mathbb{P}(E \cap F^c) \\ & + \mathbb{P}(E \cap F) \\ & + \mathbb{P}(E^c \cap F). \end{aligned} \qquad (\star)$$



Similarly, $E \cap F^c$ and $E \cap F$ are disjoint, and their union is E.

Additivity: $\mathbb{P}(E) = \mathbb{P}(E \cap F^c) + \mathbb{P}(E \cap F)$, or

(A) $\mathbb{P}(E \cap F^c) = \mathbb{P}(E) - \mathbb{P}(E \cap F)$.

Similarly, repeating argument with roles of E and F reversed,

(B) $\mathbb{P}(E^c \cap F) = \mathbb{P}(F) - \mathbb{P}(E \cap F)$.

Plug (A) and (B) back into $(\star)$ to get

$$\begin{aligned} \mathbb{P}(E \cup F) = {} & \mathbb{P}(E) - \mathbb{P}(E \cap F) \\ & + \mathbb{P}(E \cap F) \\ & + \mathbb{P}(F) - \mathbb{P}(E \cap F) \\ = {} & \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F) \end{aligned}$$

$+ \mathbb{P}(F) - \mathbb{P}(E \cap F)$

$= \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F)$. //

Examples:

Ex 1: $\Omega = \{H, T\}$ has 4 events, $\Omega = \{H, T\}$, $\{H\}$, $\{T\}$, $\emptyset$.

$\mathbb{P}(\Omega) = 1$, $\mathbb{P}(\{H\}) = \frac{1}{2} = \mathbb{P}(\{T\})$, $\mathbb{P}(\emptyset) = 0$.

Can check that $\mathbb{P}$ is a probability measure.

Alternatively, to model a "biased coin" with H probability $0 < p < 1$,

$\mathbb{P}(\Omega) = 1$, $\mathbb{P}(\{H\}) = p$, $\mathbb{P}(\{T\}) = 1 - p$, $\mathbb{P}(\emptyset) = 0$.

Ex 2: $\Omega = \{1, 2, \ldots, 6\}$, there are $2^6 = 64$ events.

Set $\mathbb{P}(k) = \mathbb{P}(\{k\}) = \frac{1}{6}$, for $k \in \Omega$.

General recipe: If Ω is finite (or countably infinite), we will assign probability to each outcome rather than each event.

- This is ok if all probabilities are nonnegative and add up to 1.
- Prob. of an event is then the sum of the prob's of individual outcomes contained in the event.
  - ↳ Mathematically: $\mathbb{P}(\{\omega_1, \omega_2, \ldots, \omega_n\}) = \mathbb{P}(\omega_1) + \mathbb{P}(\omega_2) + \ldots + \mathbb{P}(\omega_n)$.
- Note there are $2^{|\Omega|}$ events, $|\Omega| = \#$ of elements of Ω (cardinality).

Ex 3: $\Omega = \{HH, HT, TH, TT\} = \{H,T\}^2$ 2 coin flips

- Model #1: Each outcome has $1/4$ probability.

- Model #2: biased coin $0<p<1$

$$\mathbb{P}(HH) = p^2, \quad \mathbb{P}(TT) = (1-p)^2$$

$$\mathbb{P}(HT) = \mathbb{P}(TH) = p(1-p).$$

Check: $\mathbb{P}(HH) + \mathbb{P}(TH) + \mathbb{P}(HT) + \mathbb{P}(TT)$

$$= p^2 + p(1-p) + p(1-p) + (1-p)^2$$

$$= (p + (1-p))^2 = 1^2 = 1.$$

Ex 4: $\Omega = \{H,T\}^n$ n coin flips (fair coin)

$= \{\text{all strings of } H/T \text{ of length } n\}$ $\qquad \overbrace{HTHTTHT}^{n=7} \in \Omega$

There are 2^n outcomes in Ω. ($|\Omega| = 2^n$. # events is $2^{|\Omega|} = 2^{(2^n)}$.)

Each has probability $1/2^n = 2^{-n}$.

Let $E_1 =$ "first flip is H" $= \{HH\cdots HH, HH\cdots HT, \cdots, HT\cdots TT, \ldots\}$.

Clearly $\mathbb{P}(E_1) = 1/2$.

This should match the sum of probabilities of all outcomes contained in E_1.

$$\mathbb{P}(E_1) = \frac{1}{2^n} \cdot (\#\text{ outcomes in } E_1) = \frac{|E_1|}{2^n} = \frac{2^{n-1}}{2^n} = \frac{1}{2}.$$